Vol.2013-MUS-98 No.26
Vol.2013-EC-27 No.26
2013/3/16

光学文字認識による タンジブルミュージックシーケンサの提案

卯田駿介[†1]　馬場哲晃[†2]　串山久美子[†2]

近年光学画像認識技術の発展により，実物体と画像認識を組み合わせた“AR”によるアプリケーションが多く報告されている．AR に利用されるマーカには，オブジェクトの形，色の他，光学マーカと呼ばれる特殊な 2 次元マーカなどが利用されるが，これらは実物体の物理特徴や印刷パターンであるため，ユーザによって作りだすことはできない．そこで著者らは，ユーザによって手軽に AR マーカを作ることができれば，より多くのユーザにとって AR システムにおけるユーザビリティの向上につながると考えた．具体的な手法としては，光学文字認識を利用して，手書き文字を AR マーカのように扱うことを検討している．本稿ではそれらの実験段階として，実物体に印刷された文字を利用したシーケンス型電子楽器「Alphabet Sequencer」を制作した．本稿では，「Alphabet Sequencer」のシステムとその体験結果を報告する．

Proposal of Tangible Music Sequencer using Optical Character Recognition

SHUNSUKE UDA[†1]　TETSUAKI BABA[†2]　KUMIKO KUSHIYAMA[†2]

Recently, many applications by AR(Augmented Reality) have been reported by improvement in image recognition technology. The form of a two-dimensional marker or an object, the color, etc. made a system producer's and cannot make many real objects used for AR by a user. Then, authors think that if AR marker can be easily made by a user, AR system becomes familiar for more people. So, we focus on using OCR(Optical Character Recognition) to make a character as an AR marker. This time, authors try using OCR in the electronic instrument field, and made tangible music sequencer "Alphabet Sequencer" as a prototype system which can sound sequence performance. This paper reports research progress of the present condition of "Alphabet Sequencer. "

1. 研究背景

電子音源の登場により，電子楽器インタフェースは多様化している．アコースティックピアノやギター等の既存楽器を模した電子楽器から，GUI(Graphical User Interface)を演奏インタフェースに応用したもの[1]，近年では Tangible User Interface 等のように実物体指向型インタフェースを利用した事例[2]が多く報告されてきた．これらインタフェースの多様化により，演奏者によるパフォーマンス性が向上した一方で，音楽に関して専門的な技術や知識のないユーザにとっても演奏することが身近になっている．

本研究では実物体に印字されたオブジェクトをテーブルトップに配置することで演奏可能なシステムを提案するが，このようなインタラクションを提供する演奏システムはこれまで幾つか報告されてきている．

このようなシステムでは実物体を「置く」「取り去る」といった直感的な手法で演奏ができるため，マウスやキーボード等の GUI を主体とした手法よりもユーザビリティの面で優れた点が挙げられる．

そのため，このような手法は音楽知識やコンピュータ操作体験のないユーザにとって，手軽に演奏可能なインタフェースと言える．

これまで報告されてきた画像処理を利用する実物体指向電子楽器にでは，実物体の色や形状，光学マーカを利用して演奏を可能としているが，上記で述べた利点の一方で，ユーザが予め用意された実物体しか利用できず，ユーザに不必要な操作制約を設けてしまっている．

著者らはこれまで手書き文字や記号を利用して演奏可能なシステム Gocen[3]を報告しており，システム体験者から手書き記号とインタラクションにおける多くの知見を得ることができた．そこで，本稿ではこれら知見をテーブルトップ型電子楽器演奏インタフェースに応用する．ただし本稿ではその第一段階として，手書き記号ではなく，実物体に印字された文字を利用したシステムを報告する．

2. 関連研究

この節では，実物体オブジェクトを配置することによってループ演奏する電子楽器インタフェースと光学文字認識を利用した電子楽器インタフェース作品の事例を挙げる．

2.1 オブジェクト配置を利用した電子楽器インタフェース

コンピュータビジョンの普及に伴い，ウェブカメラや赤外線カメラを利用した，オブジェクト配置によってループ演奏を可能とする電子楽器インタフェースが多く提案されている．reacTable[2] は，立方体や直方体などの実物体オブジェクトをシステム専用のテーブルの上で，「置く」「回

ⓒ 2013 Information Processing Society of Japan

Vol.2013-MUS-98 No.26
Vol.2013-EC-27 No.26
2013/3/16

す」「移動する」といった操作を行うことで，演奏可能な電子楽器である．オブジェクトにはシステム専用の 2 次元コードが描かれており，この 2 次元コードをテーブルの下に設置した赤外線カメラで検出することによって，オブジェクトの座標や種類を検出している．reacTable 同様，tangible user interface for music learnig[4] もまた，2 次元コードを利用したオブジェクト配置による電子楽器である．2 次元コードを利用したこれらの電子楽器は，オブジェクトの移動や回転といった操作を伴う演奏を行いたい場合のオブジェクト認識に適し，演奏パフォーマンスの向上につながっていると言える．

LEGO STEP SEQUENCER[5] や Bubblegum Sequencer[6] は任意の領域に配置されたオブジェクトを PC に接続されたウェブカメラで取得し，取得した画像から得られる色と座標情報を元にループ演奏を行う電子楽器である．演奏動作はオブジェクトを「置く」「はずす」という単純な運動で，演奏技術がなくとも簡単に演奏をおこなうことできる点で，本研究と類似する．

また，実物体オブジェクトの形状によって音色が変わる電子楽器として，The Table is The Score[7] や Xenakis[8] などが挙げられる．

しかし，これらの研究で音色の選択に利用されている情報は，ユーザにとっては意味理解の困難な 2 次元コードや，システム制作者が準備したモノである． 本研究では，ユーザが日常生活で使っている文字を音色選択のマーカとした点が特徴的である．

2.2 光学文字認識を利用した電子楽器

文字を AR マーカとした電子楽器として，Gocen が挙げられる．Gocen はカメラの内蔵された専用デバイスを用いて，リアルタイムで五線譜の情報を読み取り，五線譜の情報に合わせて演奏を可能とする電子楽器である．Gocen では楽器，調などを，手書き文字を利用して選択することができる．

3. システム概要

本システムは 60×60×4mm の予めアルファベットの記述されたカード型オブジェクト，カード型オブジェクトを配置する台座（図 1），200 万画素のスタンド型書画カメラ，コンピュータからなる．現状のシステムは研究初期段階のプロトタイプであるため，文字認識の容易さを優先し，手書き文字ではなく，予め文字をカードに記述した．アルファベットの記述されているカードは，2mm 厚の白色アクリル材と半透明乳白色アクリル材を重ねて制作した．アルファベットの記述されている表面は 50×50mm の白色アクリル材に黒色で文字列を記述し，半透明乳白色のアクリル材で 10mm 幅の外枠を形成している．カードを配置する台座は，大きく分けて 2 つのエリアからなっており，カード 16 個分を縦 4 つ，横 4 つに並べることのできるグリッド状の右側のエリアとカード 4 つを縦 1 列に並べられる左側のエリアからなる．右側のエリアはループ演奏の一小節分を表しており，このエリアにアルファベットの記述されたカードを載せることで，どの音高をどのタイミングで鳴らすかが決まる．タイムラインは図 2 のように，左上から右下に向かって進行する．一方，左側のエリアは楽器の変更や右側エリアに置かれている情報の保持，リセットなどの操作を行うユーティリティーエリアである．

コンピュータソフトウェアは Xcode 上で作成し，openframeworks[9] ライブラリを利用し， ofxMidi，ofxopenCV，などの openframeworks 上で使用できる addon を利用した．音源は Native Instruments 社の Kontakt Player[10] を利用した．



図 1　専用の台座
Figure 1　Original Board

3.1 ループ演奏

本システムは現状では，1 ループ 16 拍で演奏するシステムとなっている．TENORI-ON[11] や PocoPoco[12] といった多くのシーケンス演奏型の電子楽器では，左側から右側に向かってタイムラインが流れ，右端までタイムラインが進行すると始めの位置にタイムラインが戻るように表現されている．これは現代に一般的な記譜法や文字列が左から右に向かって時間軸が進行していくことからもわかるように，人にとって自然な流れだと言える．そこで本システムでも図 2 のよう，左上から右下に向かってタイムラインが進行していくようデザインした．

3.2 処理の流れ

上述したとおり，openFrameworks 用に用意されている OpenCV，OCR のライブラリを利用して，上方に取り付けられたウェブカメラから画像を取得し，取得した画像の 2

ⓒ 2013 Information Processing Society of Japan

Vol.2013-MUS-98 No.26
Vol.2013-EC-27 No.26
2013/3/16

値化補正を行い，文字を認識している．一定の時間でループするタイムラインに合わせて，2 値化した画像を取得したい領域のみ抽出し，OCR ライブラリを利用して文字認識を行っている．

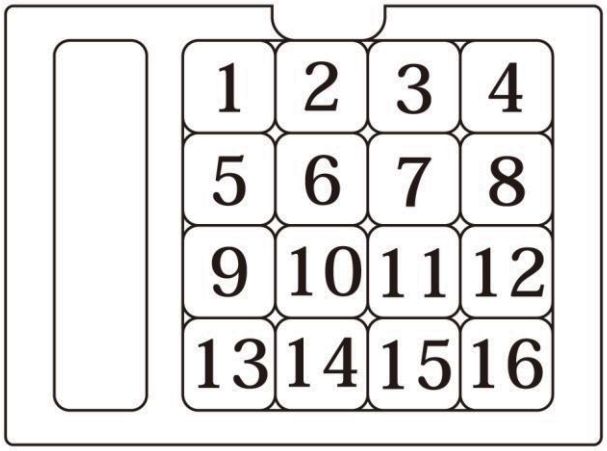


図 2 タイムラインの進行順
Figure 2　Sequence

3.3 演奏における文字の利用

本研究では従来のオブジェクト配置によるループ演奏用音楽インタフェースでは用いられなかった，文字，文字列を AR マーカとしている点が特徴である．本システムでの文字の用途を以下に 3 つ示す．

(1)音高

国ごとによって違いはあるが，古くから音高は文字によって表されてきた．例えば，イタリア音名，ドイツ音名，英語音名，フランス音名，日本音名などがある．本システムでは光学文字認識を利用する上で，凡庸性の高いアルファベットを用いた音名を利用することとした．現行のシステムでは,イタリア音名と英語音名,2 種類に対応した(図 3)．2 種類の音名を利用することで，音楽学習環境の異なるユーザでも，スムーズに利用できることを期待する．

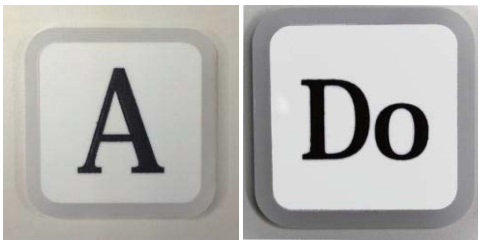


図 3　英語音名とイタリア音名の 2 種類の音階カード
Figure 3　Scale Card

(2)楽器音の選択

楽器の表記には省略記号を用いることとした．省略記号と比べ，楽器本来の名称を用いたほうが，人にとって理解しやすいと考えられるが，楽器本来の名称を用いた場合，楽器ごとに文字数にばらつきがあり，文字の認識率の低下につながると考えたためである．



図 4　楽器を示すカード
Figure 4　Instruments Card

(3)ループ中の再生情報の保持，リセット

多くのテーブルトップ型の電子楽器は，従来の古典楽器では出来なかった,「ユーザ一人で複数の楽器を同時に再生する」ことが容易となっている. 本研究でもその点を踏まえ，楽器ごとに演奏情報を保持していくレイヤーの考え方を用いて，複数の楽器を動的に演奏できるようにした．楽器ごとの演奏情報は，図 5 の「REC」，図 6 の「RESET」のカードを用いる．



図 5　演奏情報の保持に使用
Figure 5　Record and Reset playing information

4. 考察

本研究では，文字を AR マーカとすることに着目し，テーブルトップ型のループ演奏楽器の制作をした. 色や形状，2 次元コードなどの AR マーカを認識対象とした従来のテーブルトップ型の電子楽器とは異なった手法で，実装するまでに至った．このように，文字を AR マーカとし，音高の違いや楽器の変更，演奏情報の保持などといった操作を行うことができ，様々な楽器を同時に演奏することができる状態になった．

しかし，現状の本システムは，本研究の初期段階のプロトタイプということもあり，手書き文字に比べ，認識処理のしやすい活字文字をカードに予め記述する手法で実装した．今後はカードの仕様を変更し，手書き文字に対応したシステムを実装したいと考えている．また，システムの都合上，現状では演奏可能な条件が限られており，より実用的にユーザが本システムを利用するためには，課題が残っ

ⓒ 2013 Information Processing Society of Japan

た．

以下に，現状の課題と解決方法，今後の展望を詳述する．

5. 課題

5.1 ハードウェア設計の課題

現状のシステムでは，上述したとおり，上方に取り付けられたカメラを用いて，台座に置かれたカードの配置情報を取得している．そのため，台座に照射される光の輝度が均一でない場合，2 値化処理が困難になり，正しく文字を認識することができない．今後は台座に照射される光が均一になるよう，LED などを利用して，台座に当たる光の輝度を調整したいと考えている．

また，タイムラインやカード配置の記録情報の描画を現状のシステムでは PC を用いて行っている．今後はより，電子楽器としての魅力を増すために，ユーザが PC を見ずに，台座とカードのみを見て演奏できるよう，台座の構造を改善する予定である．

5.2 ソフトウェアプログラムの課題

現状のプログラムでは，異なった楽器による和音を奏でることはできる一方，同一楽器による和音を奏でることができない．ユーザによって音高を選べる本システムの利点を生かすためにも，同一楽器音による和音演奏が必要であると著者は考える. 今後, プログラムを改善していくことで，この課題を解決したいと思う．

6. 展望

(1)スマートフォンアプリへの応用

将来は，本システムをより多くのユーザに利用してもらうことを想定し，スマートフォンアプリケーションへの応用を考えている．例えば，メッセージカードをスマートフォンのカメラで読み取ると，メッセージカードに書かれた文字列に従って，音楽が流れ，ただのメッセージカードではなく，BGM つきのメッセージカードとなる．

(2)手書き文字を AR マーカとしたインタフェースの研究

現代において，パーソナルコンピュータやスマートフォンのようなキーボードを打つだけで文字を入力できる電子デバイスの普及に伴い，現代人が手書きで文字を書く機会は減少している．確かに，キーボード入力には，手軽に文字を入力できるという利点があるが, 一方で手書き文字には，人それぞれの個性があり，キーボード入力の文字には無い魅力があると著者は考える．例えば，大切な人に手紙を書く場合，キーボード入力による活字文字での手紙よりも，心をこめて，手書きで手紙を書こうと思うことでしょう．今後は手書き文字の持つ魅力を生かし, 電子楽器に限らず, 手書き文字を AR マーカとしたモノを制作したいと考えている．

参考文献

1 ） Sadam Fujioka，Osamu Sambuichi，and Shigenobu Nakamura．2006． Code: an interactive musical performance software using keywords． In *ACM SIGGRAPH 2006 Sketches* (SIGGRAPH '06). ACM， New York， NY， USA， ， Article 90 ． DOI=10. 1145/1179849. 1179962 http://doi. acm. org/10. 1145/1179849. 1179962
2) Sergi Jordà. 2010. The reactable: tangible and tabletop music performance. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '10). ACM, New York, NY, USA, 2989-2994. DOI=10.1145/1753846.1753903 http://doi.acm.org/10.1145/1753846.1753903
3 ） Tetsuaki Baba， Yuya Kikukawa， Toshiki Yoshiike， Tatsuhiko Suzuki， Rika Shoji， Kumiko Kushiyama， and Makoto Aoki， “Gocen: a handwritten notational interface for musical performance and learning music．”， In ACM SIGGRAPH 2012 Emerging Technologies (SIGGRAPH '12). ACM， New York， NY， USA， Article 9， 1 pages． DOI=10．1145/2343456．2343465 http://doi．acm．org/10．1145/2343456．2343465
4 ） E． Costanza， S．B． Shelley and J． Robinson， "INTRODUCING AUDIO D-TOUCH: A TANGIBLE USER INTERFACE FOR MUSIC COMPOSITION AND PERFORMANCE "Digital Audio Effects， 2003
5 ） Guido Lorenz， "Lego Step Sequencer"， ADVANCE HACKATHON，2010
6 ） Hannes Hesse， Andrew McDiarmid and Rosie Han， "Bubblegum Sequencer"， De-Bug Magazine Nr． 135 ，2009
7 ） Levin， G． "The Table is The Score: An Augmented-Reality Interface for Real-Time， Tangible， Spectrographic Performance．" Proceedings of the International Conference on Computer Music 2006 (ICMC'06). New Orleans， November 6-11， 2006.
8 ） Markus Bischof， Bettina Conradi， Peter Lachenmaier， Kai Linde， Max Meier， Philipp Pötzl， Elisabeth André ，"XENAKIS – Combining tangible interaction with probability-based musical composition"，Proceedings of the Second International Conference on Tangible and Embedded Interaction (TEI'08)， Feb 18-20 2008
9 ） openFrameworks community． openframeworks． www．openframeworks．cc
10 ） Native Instruments． Kontakt player． http://www．native-instruments．com/
11) ヤマハ株式会社， “TENORI-ON”
12) 金井隆晴，菊川裕也，鈴木龍彦，馬場哲晃，串山久美子，"PocoPoco: 実物体の動きを利用した楽器演奏インタフェース", 情報処理学会論文誌，一般社団法人情報処理学会，Vol. 53， No. 3， pp1050-1060， 2012

ⓒ 2013 Information Processing Society of Japan